APPARATUS AND METHODS FOR PROVIDING TELEVISION SPEECH IN A SELECTED LANGUAGE

BACKGROUND OF THE INVENTION

The present invention relates to television systems, and more particularly to

5   apparatus and methods for allowing a television program to be provided in a language

other than that recorded with the program.

Television programs include both a video portion and an audio portion. The

audio portion is recorded in a language that is typical for the locale in which the

program is broadcast. However, not all residents of a particular locale speak the same

10   language. Accordingly, it would be advantageous to provide for the selection of a

particular language in which a viewer will be able to best enjoy a particular television

program.

Prior art solutions to the language problem have generally focussed on the

provision of one or more additional audio signals, each carrying the audio portion of

15   the television program in a different language. For example, various proposals for

digital television transmission include a provision for a second audio program (SAP)

which can be used to provide, e.g., television audio in a second language. A problem

with such a solution is that each separate audio signal requires additional bandwidth

in the broadcast signal. The use of such additional bandwidth is undesirable, as it

20   consumes space that could otherwise be used for revenue generating services, such as

additional programming.

In the past, closed caption data has been provided to enable the hearing impaired to view the audio portion of a television program as text. Such data is carried in analog and digital television signals in accordance with applicable television standards, such as the National Television Systems Committee (NTSC) standard for

5   analog television in the United States, and the Moving Picture Experts Group (MPEG) standards for digital television. In the past, closed caption data has only been used for such display of text.

It would be advantageous to provide a system for enabling a viewer to choose any one of a number of different languages for the audio portion of a television program.

10   It would be further advantageous for such a system to provide different languages without requiring additional bandwidth for each language.

The present invention provides a television audio system having the above and other advantages.

SUMMARY OF THE INVENTION

The present invention enables a television viewer to select the language in which television speech will be provided. In order to provide this ability, closed caption data is extracted from the television signal. The closed caption data is

5   representative of words. The extracted closed caption data is processed in a speech synthesizer to provide the words as speech in the desired language.

A user interface is provided to enable the user to select one of a plurality of languages capable of being provided by the speech synthesizer. The user interface can include, e.g., a television on-screen display. In such an embodiment, the user

10   interacts with the on-screen display via a television remote control.

Since the television signal will typically already include an audio portion in a first language, this audio portion will be muted if another language is selected. In this manner, the audio portion carried with the television program will not interfere with the audio output of the speech synthesizer.

15   In one embodiment, the closed caption data is first converted to text. The text is then converted to speech. The closed caption data can be representative of words in the desired language. Alternatively, the closed caption data can be representative of words in a language that is different from the desired language, in which case processing will be provided to translate the words into the desired language prior to

20   synthesizing speech therefrom.

Apparatus for implementing a preferred embodiment of the invention includes a closed caption processor adapted to extract closed caption data from a television signal having an audio portion in a first language, the closed caption data being representative of words. A speech synthesizer is provided to convert the words

25   represented by the closed caption data to speech in a second language.

The user interface, which enables user selection of the second language, can comprise, for example, a remote control that allows the user to interact with a television on-screen display. A mute circuit is provided for muting an audio portion of the television signal when replacement speech is provided from the speech

5   synthesizer.

The invention can also be implemented, at least in part, in a software program adapted to provide television speech in a selected language. Such software can include a closed caption processor module adapted to extract closed caption data from a television signal having an audio portion in a first language, said closed caption data

10   being representative of words. The software can further include a speech synthesis module adapted to convert the words represented by said closed caption data to speech in a second language.

The software program can further comprise a user interface module for enabling a user to select one of a plurality of different languages as the second

15   language. The user interface module can, for example, include software code for generating an on-screen display to enable the user to select the desired second language using a remote control. A mute module can also be provided for actuating a mute circuit to mute an audio portion of the television signal when replacement speech is provided from the speech synthesis module.

20   The closed caption module of the software program can be designed to convert the closed caption data to text for processing into speech by the speech synthesis module. The text can be provided in the second language. Alternatively, the text can be in a language other than the selected second language, in which case the speech synthesis module can be adapted to translate the text to the second

25   language for processing into speech. The software program can be provided on a machine readable media.

A method is also disclosed for providing audio from a television signal in a selected one of a plurality of different languages, where the television signal includes

the audio in one of the languages. A user selects one of the languages. If the selected

language is not the language included in the television signal, the language included

in the television signal is converted to the selected language for audio presentation to

the user. In one implementation, the language is converted from text provided in a

5   closed caption signal. In another implementation, the language is converted from the

audio portion of the television signal.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 is a block diagram showing the main components of a system in accordance with the present invention; and

Fig. 2 is a block diagram showing an example software implementation of the invention.

DETAILED DESCRIPTION OF THE INVENTION

The present invention uses closed caption data representative of words, in

conjunction with a speech synthesizer, to provide television audio output in a desired

language. In this manner, the television viewing experience is enhanced by allowing

5   a viewer to select a language other than the main language associated with the

program, as the language that the user will hear when listening to the program. In the

past, when a viewer wanted to listen to a program in a language other than the

language associated therewith, the content provider would have to supply a second

language with the program. This requirement limited the number of languages

10   available, and placed the burden on the content provider to supply additional

languages. The present invention overcomes this problem by utilizing the closed

caption data and a text-to-speech converter (i.e., a "speech synthesizer") to convert

the closed caption text to a user selected language. The selected language is then

presented to the user instead of the main language carried by the program.

15   Figure 1 illustrates the relevant hardware components of the invention. A closed

caption processor 10 extracts closed captioning data (e.g., in the form of text) from a

received television program. The closed captioning data is provided to a text-to-

speech processor 12, which includes text recognition and/or translation software for

converting the closed captioning data to a selected language. Although Figure 1

20   illustrates the capability of the processor 12 to convert the closed caption text from,

e.g., English to Spanish, German, French or Russian, it should be appreciated that any starting language can be accommodated and any ending language can be provided by providing appropriate software.

Text-to-speech processors are well known in the art, and any suitable such device

5   can be used in order to implement the present invention. For example, Oki Electric Industry Co., Ltd. of Tokyo, Japan markets its model MSM7630 multi-lingual speech control processor (SCP) with text-to-speech synthesis capability in six languages including American English, European English, French, German, Spanish, and Japanese. This product uses a single large scale integrated circuit chip with a 12-bit

10   D/A (digital-to-analog) converter to provide a natural sounding voice using time domain - pitch synchronous overlap-add technology to replicate waveforms in human voices. Both parallel and serial interfaces are provided to accommodate various implementations. A user dictionary can be programmed to expand vocabulary, and is available in Flash-ROM (read only memory) for easy upgrades.

15   The text-to-speech processor 12 of the present invention is programmed to provide as output any desired one of a number of selectable languages. The languages can be changed and/or expanded, for example, by providing additional software modules that are either downloaded to the device, or installed by inserting a non-volatile memory card (e.g., Flash-ROM) or the like into a receptacle in the

20   device. A user can be provided with an electromechanical switch, or with a graphical user interface (GUI) or the like in order to make the language selection. In a

preferred embodiment, a GUI is provided on the user's television screen using, e.g.,

standard on-screen-display (OSD) hardware and software 18, which displays a list of

available languages that the device is capable of "speaking." The user can then select

a language using the television remote control 14, for example, by pressing a button

5   (such as a number button) thereon that corresponds to the desired language. The

remote control response is detected by a user interface 16 (e.g., via infrared (IR)

signal reception), which actuates the text-to-speech processor to convert the received

closed caption text to the requested language.

When a language other than the main language in which the program is received

10   is selected, the text-to-speech processor 12 provides a switching signal to a switch 20,

in order to couple the output of the text-to-speech processor to the television audio

amplifier 22 and speaker 24. When the switch 20 is coupled to the text-to-speech

processor, the original program audio is muted, as it is disconnected from the audio

circuitry 22, 24. When it is desired to hear the original program language, the switch

15   20 is switched to couple the original television audio output to the amplifier 22 and

speaker 24.

Figure 2 provides a flowchart of processing and software components that can be

used to implement the invention. In particular, user input 30 (i.e., language selection)

is provided to a processor 32, which can be the microprocessor already provided in a

20   television settop. An example of a microprocessor controlled settop box is the DCT-

5000 manufactured by the Broadband Communications Sector of Motorola, Inc.,

Horsham, Pennsylvania, USA. The processor also receives a digital television signal,

which contains a main language audio portion as well as closed caption data. It is

noted that although Figure 2 illustrates the processing of a digital television signal,

closed caption data is also carried in analog television signals, and can be extracted

5   for input to processor 32 in digital form.

The processor 32 provides television video 34 and audio 36 to a user's television

in a conventional manner. In accordance with the present invention, software 38 is

included for use in providing the television audio 36 in a selected alternate language.

The software 38 can reside in a non-volatile memory portion of the settop, such as in

10   ROM, and can be installed at the factory or warehouse, or downloaded into the settop

via the cable television network, via telephone lines, or via a wireless communication

path, for example. Alternatively, the software can be stored in a hard drive or other

memory portion of a personal versatile recorder (PVR) device, personal computer

(PC) attached to the settop, or the like.

15   As indicated in Figure 2, the software 38 includes a module for implementing the

closed caption processor which extracts the closed caption (CC) data from the

television signal. The closed caption processor module provides the closed caption

data in text form to a speech synthesis module, which translates the text to the desired

language, and provides the translated text as speech to the audio circuits of the user's

20   television or other video appliance, such as a video tape recorder, PVR, or the like.

Software 38 also includes a user interface module, which provides an on-screen

display for enabling users to select the language which they want to hear. The interface module also handles the decoding of user input signals from the television (or settop, VCR, PVR, etc.) remote control. A mute module is also provided to mute the main program audio output so that the selected alternate language can be heard

5   via the television audio system. It should be appreciated that the implementation shown in Figure 2 is for purposes of illustration only, and that other implementations can be provided in accordance with the invention.

It should now be appreciated that the present invention provides a new use for closed caption data. Instead of using such data to present text to the hearing

10   impaired, it is used to provide audio speech in different languages to viewers who can hear the speech. As an alternative, the closed caption text can be carried in the television signal in different languages, which can be directly input into a text-to-speech processor for conversion to speech without any need for translation.

Although the invention has been described in connection with a specific

15   embodiment thereof, it should be appreciated that various modifications and adaptations can be made thereto without departing from the scope of the invention, as set forth in the claims.